

REMARKS

This application contains claims 1-24. Claims 4-7 and 9-24 have been canceled without prejudice. Claims 1 and 8 have been amended. No new matter has been introduced. Reconsideration is respectfully requested.

Applicant is not conceding that the subject matter encompassed by claims 1 or 4-24, prior to this amendment is not patentable over the art cited by the Examiner. Claims 1 and 8 were amended and claims 4-7 and 9-24 were canceled in this amendment solely to facilitate expeditious prosecution of the subject matter of claims 1-3 and 8. Applicant respectfully reserves the right to pursue claims, including the subject matter encompassed by the claims as presented prior to this amendment and additional claims, in one or more continuing applications.

Claims 1-22 were rejected under 35 U.S.C. 101 for failing to recite a functional interrelationship with a computer. While disagreeing with the grounds of rejection, Applicant has amended claim 1, for the sake of expediting prosecution, to recite that the steps of the claimed method are carried out by a computer. The amendment is literally supported in the specification in paragraph 0027 *et seq.* (Paragraph numbers in these remarks refer to the published version of this application, US 2005/0123203.) In view of this amendment, claims 1-3 and 8 are believed to meet the requirements of 35 U.S.C. 101.

In view of the cancellation of claims 4-7 and 9-24, the rejection of these claims under 35 U.S.C. 101 is moot.

Claims 1-22 were rejected under 35 U.S.C. 112, first paragraph, for lack of enablement of certain claim terms. Applicant has amended claim 1 in order to clarify the relationship between the claim terms and the specification of this application as filed.

In relation to claim 1, the Examiner stated that the present patent application failed to teach "character codes that are likely to be generated due to a segmentation error." Applicant questions this finding, since paragraphs 0006 and 0008 of the specification both demonstrate explicitly how certain character codes may sometimes occur due to incorrect segmentation. Nevertheless, for the sake of expediting prosecution, Applicant has amended claims 1, 9 and 17 to refer to "character codes that are known to result frequently due a segmentation error." This claim language is literally supported and explained in paragraph 0031.

In relation to claim 2, the Examiner stated that the present patent application fails to teach "assigning a respective eight-bit code that is not included in the predetermined set to replace each of the sequences." Applicant respectfully traverses this rejection. Paragraph 0032 explains clearly how ASCII codes that are not used in OCR output (such as codes corresponding to Greek letters) may be assigned to extension characters. It is well known in the art that ASCII codes are eight-bit codes. Therefore, claim 2 is clearly enabled by the specification.

Claims 1-22 were rejected under 35 U.S.C. 112, second paragraph, for lack of clarity in use of the term "likely." As noted above, Applicant has amended claim 1 to remove this term and recite instead "character codes that are known to result frequently due a segmentation error." In view of this amendment, all of the claims still pending in this application are now believed to meet the requirements of 35 U.S.C. 112.

In view of the cancellation of claims 4-7 and 9-24, the rejection of these claims under 35 U.S.C. 112 is moot.

Claims 1, 2, 5, 8-10, 12, 13, 16-18, 20, 21 and 24 were rejected under 35 U.S.C. 102(b) over Kagehiro et al.

(U.S. Patent 6,246,794). Applicant has amended claim 1 in order to clarify the distinction of the claimed invention over the cited art. Claim 1, as amended, incorporates the limitations of claims 4-7, which have been canceled. Claim 8 has been amended for proper dependence in view of the cancellation of claim 4. Claims 9-24 have been canceled without prejudice.

Claim 1, as amended, recites a computer-implemented method for encoding characters in an area of image, in which certain sequences of character codes (referred to in the specification, paragraph 0031, as "critical combinations") are identified as resulting frequently from incorrect segmentation of original characters. These critical combinations are assigned respective extension character codes. After segmenting the image, a pattern recognition process is applied to each segment in order to generate an input string of character codes. A modified string is then generated by replacing at least one of the critical combinations of character codes in the input string with the corresponding extension character code. For example, referring to paragraph 0032 in the specification, occurrences of the critical combination "rn" (which is known to result from incorrect segmentation of the original character "m") are replaced at this stage by an extension code, which in this case is the ASCII code for the Greek character "α".

To determine the output string, the computer computes "edit distances" between the modified string and known strings in a directory. The distances are based on costs of edit operations (changing one symbol into another single symbol, deleting a symbol from a string, or inserting a symbol into a string, as defined in paragraph 0003). For edit operations involving at least some of the extension character codes, however, a cost of zero is assigned to transformation of each such code into its corresponding original character code. Thus,

continuing the above example, the cost of transforming "α" back into the original "m" is zero. This innovation permits segmentation errors to be corrected simply and reliably, while using existing pattern recognition and string matching processes without substantial modification.

Kagehiro describes a method of reading characters that uses different segmentation hypotheses to deal with types of segmentation errors (col. 15, lines 3-62, illustrated in Figs. 20-23). Kagehiro relates to these segmentation errors in the course of feature extraction, i.e., as part of the process of identifying character patterns. By contrast, as explained above, the method of claim 1 deals with possible segmentation errors by operating on character codes after the pattern recognition process has been carried out, rather than on the image features themselves.

Furthermore, Kagehiro does not relate at all to "edit distances," in the sense in which this term is used in the art and is explicitly defined in the present patent application (paragraph 0003, cited above). Although the Examiner pointed out, in reference to claim 5, that Kagehiro uses a distance between nodes in recognizing character patterns (col. 9, lines 20-21), this is an actual physical distance between elements of a graphical representation of a pattern. Kagehiro neither teaches nor suggests the use of edit distances for any purpose. The Examiner also acknowledged, in reference to claims 6 and 7, that Kagehiro does not relate to costs of edit operations or to assigning a zero cost to certain operations, as is recited in amended claim 1.

Andersen et al. (U.S. Patent 7,240,062, cited by the Examiner against certain dependent claims) describes a method for creating a searchable word index using multiple recognition engines to provide different interpretations of a word in a scanned document

(abstract). Contrary to the Examiner's assertion, Andersen does not teach or suggest the use of either edit distances or costs of edit operations. Andersen mentions "cost" only in the context of a sequence of words to be recognized in a document ("The maximal cost," as illustrated in Fig. 3). This chance use of a word that appears in the claims of the present patent application cannot possibly be taken to teach anything about how costs might be assigned to edit operations or used in correcting segmentation errors in encoded characters. Andersen does not even mention zero costs, and it goes without saying that he cannot be taken to suggest assigning a zero cost to certain operations as recited in claim 1.

Therefore, claim 1, as amended is patentable over Kagehiro, whether Kagehiro is taken alone or in combination with Andersen. In view of the patentability of claim 1, dependent claims 2 and 8 are also believed to be patentable.

Claims 3, 6, 7, 11, 14, 15, 19, 22 and 23 were rejected under 35 U.S.C. 103(a) over Kagehiro in view of Anderson. Claims 6, 7, 11, 14, 15, 19, 22 and 23 have been canceled. In view of the patentability of claim 1, dependent claim 3 is also believed to be patentable.

Applicant believes the amendments and remarks presented hereinabove to be fully responsive to all of the grounds of rejection raised by the Examiner. In view of these amendments and remarks, Applicant respectfully submits that all of the claims in the present application are in order for allowance. Notice to this effect is hereby requested.

Please charge any fees associated with this response
to Deposit Account 09-0468.

Respectfully submitted,

By: _/Suzanne Erez/_____

Suzanne Erez

Reg. No. 46,688

Phone No. (972) 4-829-6069

Date: 6 March 2008

IBM Corporation

Intellectual Property Law Dept.

P. O. Box 218

Yorktown Heights, New York 10598